Evidence for Observation
of Virtual Radio Cherenkov Fields

Alice Bean, John P. Ralston and James Snow

Department of Physics and Astronomy, University of Kansas, Lawrence KS 66045

Abstract

We present evidence for observation of virtual electromagnetic fields in the radio domain from experiment T926 at the Fermilab Meson Test Beam Facility. Relativistic protons with 120 GeV energy traversed a sealed electromagnetic cavity and were observed in the radio regime of 200MHz-GHz. Closely related to ordinary Cherenkov radiation, which we also measured, the virtual fields require no acceleration for their existence. The experiment is also the first observation of fields from hadronic showers, an independent and new confirmation of coherent radio emission from ultra-relativistic particles. Conditions of very low signal to noise were overcome by a novel and unbiased filtering strategy that exploits exhaustive studies of correlations in the noise backgrounds. Linear scaling of the signal region with the number of beam particles provides evidence of coherence. Extrapolation to measurement of the field of a single relativistic proton charge is consistent within errors. Our study also illustrates new data processing methods that may be applied broadly in conditions of extremely low signal to noise.

Key words:
PACS: 29.40.Ka, 41.60.Bq, 95.55.Vj , 14.70.Bh

1 Introduction

We conducted experiment T926 at the Fermilab testbeam facility to measure electromagnetic fields of relativistic bunches of protons passing close to radio antennas. The fields of moving charges were studied both in free space (virtual fields) as well as with radio antennas embedded in wax (Cherenkov fields). We present evidence for observation of the virtual electromagnetic fields from the relativistic protons traveling in air. We compare these measurements with detection of Cherenkov radiation, where particles move at about 150% of light speed in the medium. Measurements made in the radio frequency regime find the real and virtual signals to be comparable. In

arXiv:1008.0029v1 [physics.ins-det] 30 Jul 2010

particular we find no evidence for dramatic differences between the "on-shell" Cherenkov fields and the virtual fields. Besides the conceptual interest of clarifying basic physics, our study also highlights new signal filtering strategies applicable to many circumstances. The signals are exceedingly small; combining standard averaging methods with a novel and highly efficient filtering strategy yields the observation.

Radio frequency (RF) signals from ultra-relativistic particle showers are the leading technology for detecting ultra-high energy cosmic ray neutrinos. RICE[1] is the prototype ice-target radio-neutrino telescope. RICE operates at the South Pole and has established the world's tightest bounds on neutrino fluxes above 10^{18} eV. ANITA[2] is a balloon-born instrument with the same purpose. ICECUBE[3] is the km^3 neutrino detection experiment underway at the South Pole currently developing a radio detection component AURA. Previous relativistic RF experiments have measured *coherent radio Cherenkov radiation* (the "Askarian effect" [4]) using electron beams with approximately 10^8 times more charged particles than available to us. We initially proposed[5] to use the full Fermilab beam towards calibration of neutrino signals. Here we describe a much more difficult experiment given the features of the facility available. To analyze the data we developed signal processing methods that could also ameliorate anthropogenic noise affecting RICE, ANITA, ICECUBE, and numerous other experiments. The technique amounts to using noise correlations that are not ideally random against noise itself.

1.0.1 *Conceptual Background*

Cherenkov radiation is a familiar tool of high energy and nuclear physics. By far the most common use of Cherenkov radiation comes at optical frequencies. The distance from source to detector is normally many millions of wavelengths, and practically indistinguishable from light of a source at "infinity." The experiment we will describe explores interesting conceptual issues of Cherenkov fields and what is meant in physics by the term "radiation."

"Radiation" is conventionally defined by fields caused by acceleration of charges, that fall like the inverse distance from the source, and that move at the speed of light. There is a certain arbitrariness in these criteria. Indeed the distinction of virtual and Cherenkov fields is partly one of terminology. One defining feature of "virtual" fields is that they cannot propagate independently to infinity. The virtual fields we explore are the electric field from moving charges, which therefore need not move at the speed of light. These fields obviously do not require acceleration of charges for their existence. Virtual fields are commonly associated with ultra-short distances and

quantum fluctuations, but their existence is much more general. The key to measuring virtual fields consists of controlling the environment around the moving charges and working in a low frequency regime where virtual fields extend to macroscopic distances. Ordinary radio-frequency instrumentation suffices to measure the virtual fields. Despite the simplicity of the situation we have not found a reference for measurement of the impulsive fields from uniformly moving charges under free-space conditions.

The experiment we will describe involves a straightforward measurement of fields moving in free space at about 99.996 % of light speed, compared to fields moving in wax at about 150% of light speed in the medium. Common understanding of Cherenkov physics might lead one to expect qualitatively different behavior, but that understanding hinges on "real radiation" propagating to infinity. Theory suggests the virtual fields should not be very different. Besides the inherent interest, our experiment sought to measure electromagnetic fields in the radio frequency regime from hadronic showers, and is the first to do so. Except for scaling due to particle numbers, the virtual fields of hadronic showers are also not expected to be dramatically different from fields or protons in free space.

The concepts require going beyond the plane wave and dipole approximations of textbook radiation theory. To develop more general solutions let $A^\mu(\vec{x}, t)$ be the vector potential in a Lorentz gauge. Seek a configuration translating uniformly along the z axis with arbitrary speed v:

$$A^\mu(\vec{x}, t) = A^\mu(\vec{x}_T)e^{i\omega(z-vt)/v}.$$

Apply the wave equation to find

$$\nabla_T^2 A^\mu(\vec{x}_T) + \frac{\omega^2}{c_m^2}(1 - \frac{c^2}{v^2})A^\mu(\vec{x}_T) = 0. \qquad (1)$$

Here ∇_T^2 is the Laplacian for transverse coordinates $\vec{x}_T$, and c_m the speed of light in a medium. A plane wave ansatz replaces $\nabla_T^2 \rightarrow -\vec{k}_T^2$ giving $\omega^2 + v^2\gamma_m^2\vec{k}_T^2 = 0$, with $\gamma_m = 1/\sqrt{1 - v^2/c_m^2}$. An on-shell solution requires $v > c_m$, from which trigonometry yields asymptotic plane waves that move away at angle $cos\theta_c = c_m/v < 1$ at speed c_m. The on-shell case reproduces Cherenkov radiation (actually predicted first by Heaviside) [6], where in a medium $c_m = c/\sqrt{\epsilon_\omega\mu_\omega}$, with dielectric constant ϵ_ω. Such fields satisfy the classic meaning of *radiation*.

Our main focus is *subluminal* velocities, $v < c$. The relevant wave packets are *not* reducible to plane waves. "On-shell" solutions with real ω, $\vec{k}_T$ *do not*

3

exist, so that the actual solutions are virtual fields. Cylindrically symmetric solutions to Eq. 1 are

$$E_\omega^z = \frac{-2iq\omega F_\omega e^{i\omega z/v}}{\epsilon v^2 \gamma^2} \left[K_0\left(\frac{\omega|\vec{x}_T|}{v\gamma}\right) - K_0\left(\frac{\omega R}{v\gamma}\right) \frac{I_0\left(\frac{\omega|\vec{x}_T|}{v\gamma}\right)}{I_0\left(\frac{\omega R}{v\gamma}\right)} \right],$$

(2)

where $\vec{E}_\omega$ is the electric field in the frequency domain and F_ω is the form factor from longitudinal charge distribution. Here K_0 and I_0 are modified Bessel functions. The singularity of the K_0 term as $x_T \to 0$ represents a point charge q. Parameter R implements boundary conditions controlling $E_\omega^z(R)$. The case $R \to \infty$, an infinite homogeneous medium, has exponentially damped solutions going like $exp(-\omega|\vec{x}_T|/v\gamma_m)/\sqrt{2\pi\omega|\vec{x}_T|/v\gamma_m}$. Otherwise the I_0 term accounts for the interior solution for cylindrically conducting walls that enforce the boundary conditions, also known as image charge effects. We arranged experimental geometry so that the $I(0)$ term is negligible, and the fields are essentially those of free space. At the same time we used a conducting cavity to shield external noise.

In the impulse approximation we measure these fields of uniform motion. From causality the tiny effects of energy conservation must be accounted for *after* the event, during a time set by the inverse frequencies involved. The transverse extent of the virtual fields Δx_T, normally assumed microscopic, may extend sideways into the macroscopic domain. Inspecting Eq. 2:

$$\Delta x_T \sim 3 \times 10^{-4} cm \, \frac{v}{c} \frac{\gamma_m}{100} \frac{eV}{\hbar\omega}.$$

(3)

Choosing $\omega \sim GHz$ in the radio frequency domain and $\gamma >> 1$ makes Δx_T macroscopic in reach. Given the frequency dependence of Eq. 3, radio antennas become the detection tool of choice. Notice that this radiation is present *en vacuo* and (if technology permitted) detectable "at walking speed".

1.0.2   Air Versus Wax

A brief examination of Eqs. 1 and 2 shows that the medium enters in terms of the dielectric constant ϵ and the effective boost parameter γ_m. Its definition is

$$\gamma_m = \frac{1}{\sqrt{1 - v^2/c_m^2}}.$$

Charges moving faster than light speed in the medium have $v > c_m$ and $\gamma_m \to i\,|\gamma_m|$ becoming imaginary, assuming ϵ is approximately real. This transformation is exactly like the forbidden case of imaginary γ often said to be absurd in free space, and cited as prohibiting particles with $v > c$. The appearance of imaginary γ_m is not absurd when used in the field expressions, and the imaginary phase simply represents energy loss to the medium.

In wax at radio frequencies the refractive index is about 1.5. The value of γ_m is a frequency-dependent complex function with magnitude of order unity. The near-field numerical repercussions of $v > c_m$ (Cherenkov fields in wax) versus $v \lesssim c$ (virtual fields in air) turns out to be rather modest, a change of "relative order one."

We found it interesting and somewhat counter-intuitive for short distance Cherenkov fields in wax and virtual fields in air *not to be extremely different*. A simple physical picture explains the physics. Cherenkov radiation is invariably pictured as a "shock-wave" seen asymptotically far from the source. In that region waves constructively interfere along the Cherenkov cone. This is the regime where photons evolve to configurations called "real", with each frequency ω eventually moving at speed $c_m(\omega)$ in directions along the normal to the expanding cone. Eventually the conditions of "radiation" are fulfilled, including energy transport away from the center and a $1/r$ scaling of amplitude with distance r.

In the zone close to the detectors, however, there is no shock wave. The phase relations of different frequencies order the fields so that those in air and wax are not very much different. An analogy with the wake of a moving boat is quite accurate. The wake is attached to the boat, and moves longitudinally with the boat at whatever speed the boat moves. The virtual field wake for charges moving at 99.996 % c in air is just the boosted Coulomb field, a kinematic consequence of motion. An ideal, lossless wake reconstructs itself coherently and does not transport any energy to infinity. Compare super-luminal motion of charges in wax. The Cherenkov field *continues to move longitudinally at the basic charge's speed of 99.996 % c*, and near the charges is comparable in amplitude to the case of air. However the field lines are drawn back by the source outstripping the propagation speed, setting themselves into shape later to add coherently on the cone, long after the charge has passed. Just as an airplane does not hit a brick-wall when its speed becomes supersonic [1], but makes a big "boom" far away, the transition from virtual to real radiation close to the source is undramatic.

---

[1] Supersonic shock wave modeling for airplanes involves a non-linear component, but the primary effect is linear just as in electrodynamics.

It was unnecessary, and we did not attempt to discriminate between the beam speed of 99.996 % c and $c = 1$, a task well beyond our timing resolution. The experiment we will describe was also not intended to develop fine resolution of overall normalization constants, and our expectations included seeing Cherenkov signals and virtual signals of about the same magnitude. Normalizations are difficult because the simulation (described below) folds together a number of factors that include amplifier gains, antenna response, and beam-antenna separation. We planned an experiment where it would be sufficient to see the fields "attached" to the proton beams to verify detection.

1.1   Experimental Overview

To measure the fields associated with moving charges we constructed a specialized apparatus ("tank", Fig. 1) at the recently established Fermilab Meson Test Beam facility[7]. The facility culls 120 GeV protons from the Main Injector and steers them towards the fixed target line. Beam intensity is limited using a collimator, followed by 300m of free flight outside a beampipe to the testing area. About once a minute, several Main Injector turns (typically 8-11) are built up and extracted in "fastspill" (maximum intensity) mode of 7-11 RF buckets, separated by 18.9 ns, a 53 MHz repetition rate. The first bucket contained $N_p \sim$ 300-600 protons, with subsequent buckets being comparable and varying slightly from spill to spill. Our experiment was the first run at the facility.

The detected signal in volts $V(t)$ is the convolution

$$V(t) = \int d\omega \, \vec{E}_\omega \cdot \vec{T}_\omega e^{-i\omega t}.$$

Here $\vec{T}_\omega$ is the vector valued, frequency-dependent transfer function from antennas, cables, amplifiers, and filters. From Eq. 2 the time scale of passage of a 100 GeV proton at 1 cm distance is of order $10^{-11}s$. The Fourier domain pulse is basically flat up to 1 THz, appearing to be a delta function spike per proton. The pulses add coherently, convolved with the time-structure of the beam, discussed shortly.

The tank apparatus (Fig. 1) enforces boundary conditions of cylindrical symmetry about the beam trajectory. Early plans[5] called for a large open volume to measure the transverse field behavior. To exclude a significant RF background at 53 MHz we adopted a thick-walled aluminum tube of diameter 48 cm, putting the 53 MHz noise well below the minimum (cutoff) oscillation mode of 154 MHz. Two 1 inch long brass biconical antennas are oriented transverse to the beam and positioned about 1 inch from the cen-
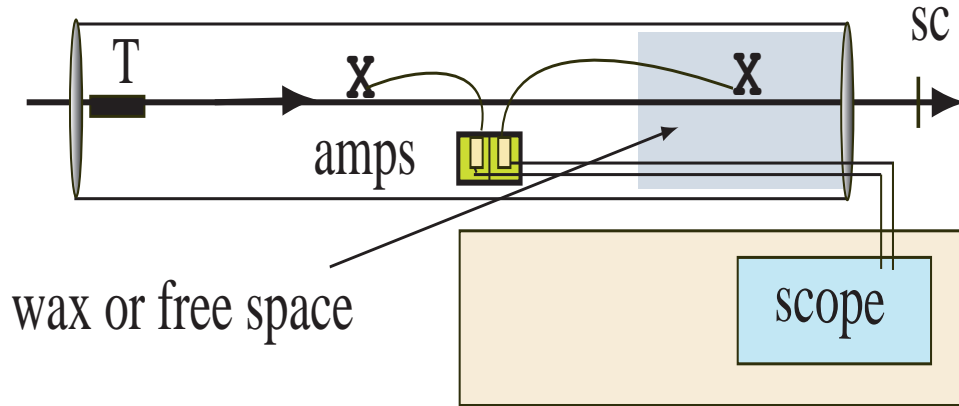
Fig. 1. The tank apparatus in schematic. The beam enters on axis of a conducting cylindrical cavity 15 feet long containing two antennas (*X*) near the axis. The trigger scintillator (sc) is 15 feet downstream. A "medium" of free space ("air") or wax fills the back 3 feet of the tank; a lead pre-shower target (T) was available in the front.

terline, separated longitudinally by 10 feet, for redundancy and the possibility of measuring causally correlated signals. A 10 foot long quad shielded coax cable connected each antenna to its own 200 MHz high pass filter and 50dB amplifier contained in a shielded box within the tank. Another 40 foot long shielded cable brought each amplified signal to our TEK7104 1GHz analog bandwidth oscilloscope in the test beam control room. Data acquisition was triggered by a 2"x2" scintillator centered on the beam and positioned 15 feet *downstream* of the tank. Thus RF noise from the phototube could not appear in the antennas until after the beam passed. The trigger RF spectrum was also measured to be in the regime below 200 MHz. Our data recorded simultaneous readout of 3 channels (2 antennas and the trigger output) with 2500 time points sampled at Δt=0.16ns intervals for a total event record of 400ns. Cable delays were chosen to put the scintillator trigger time near the expected first antenna hit and near the center of the event record. This placement retained an event-by-event "noise region" in the first part of each record, well separated by causality from the "signal region". Extensive collection of noise was a key element of the experiment.
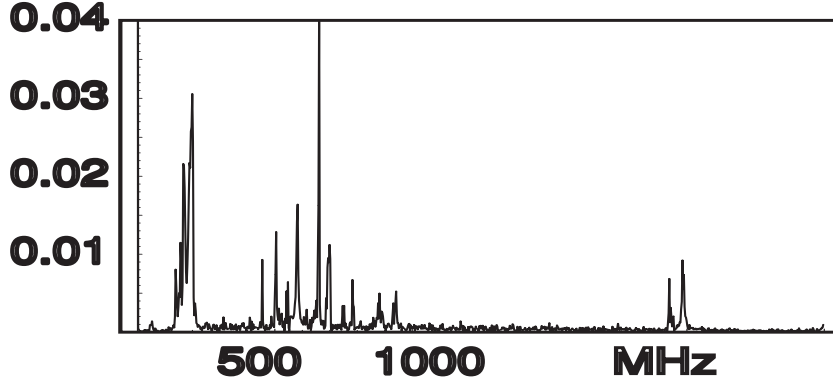
Fig. 2. Fourier absolute spectrum (absolute value of Fourier transform) of the average of 400 noise segments in channel 1 in arbitrary units. The peaks above 420 MHz are tank resonances. The structure around 200 MHz is a mode below the nominal cutoff frequency.

1.1.1 RF Studies

Despite heavy shielding, noise of manmade origin (*e.g.* cell phone bands) dominated. We installed metal endcaps on both ends of the tank. A 1 inch diameter hole in the front cap was covered with 0.040″ thick aluminum to shield radiation while minimizing beam interactions. Calculations solved the boundary-value problem for our geometry to see if transition radiation would be an issue. Unlike the usual case for optical frequencies in dielectrics, transition radiation is negligible in our circumstances. Peaks in the frequency spectrum calculated for the cavity were found to be in good agreement with lines observed in the noise (Fig. 2). In particular the "fundamental" modes dominate the central region. Identifying every single line is difficult since each mode of given transverse quantum numbers has numerous submodes of longitudinal oscillations. A discrepancy with expectations occurs in the region of about 200 MHz and below the nominal tank cutoff frequency. The explanation is leakage of exponentially damped modes (complex longitudinal wave number k_z). Sample to sample variations in the amplitudes were large. They encode background conditions of myriad origins, which are *not* described by textbook assumptions of "uncorrelated" noise.

We pre-calibrated the system's transfer function in the lab with a pulse generator. This left uncertainties of order a few dB from the coupling to the modes in the tank and its local environment. System response, denoted T_ω^{click}, was remeasured directly from the "click" spark of a piezoelectric cigarette lighter, which is close to a delta-function in time, calibrating the tank-antenna-amplifier-filter-cable assembly in one step. The time struc-
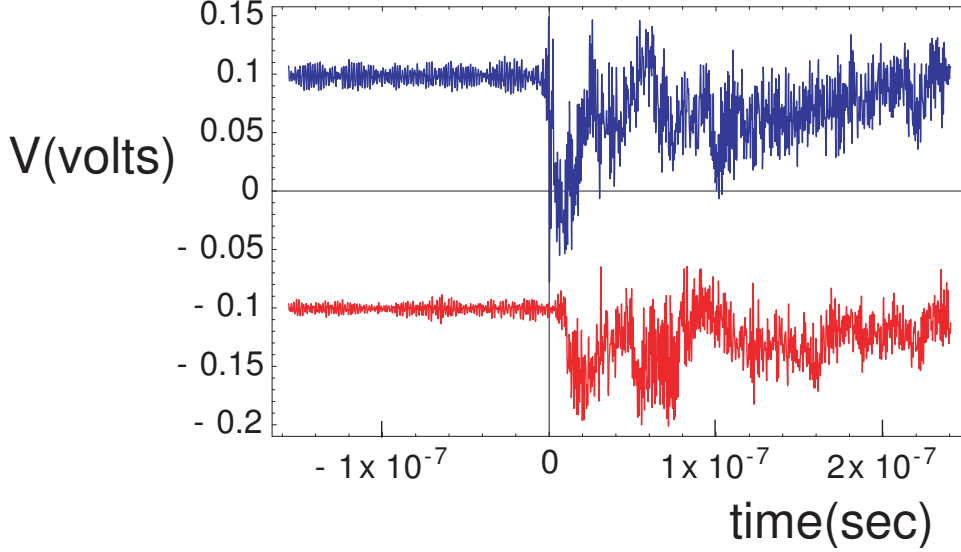
8

Fig. 3. Average antenna response over 138 runs of "click files" made with a piezo-electric spark to calibrate the system. Top and bottom are channels 1, 2 (upstream, downstream) with timing offset showing light-speed separation of the antennas. The vertical black line shows the trigger location. Units on the horizontal axis are seconds, vertical axis volts, in arbitrary scale.

ture of response is directly shown to be much shorter than the beam bucket structure. Fig. 3 shows the average of 138 click-runs. In these and other figures the vertical scale (units of volts) depends on amplifier gains and cable attenuation and so is plotted in arbitrary scale. We naturally use the same scale in passages about quantitative comparison.

The "click runs" also verified the expected time difference of a causal signal between the two antenna channels, approximately 10ns (Fig. 3). Thus our calibration, and subsequent data processing, are based on truly radiative processes.

1.1.2 *Target, Beam and Background Noise*

The test-beam cycle of one spill per minute and a fixed total running time dominated our total number of events. We collected two datasets of approximately 400 events, treated independently in the analysis. *Set-A* used antennas in an empty tank ("air") to look for virtual radiation. In *Set-W* the downstream antenna was embedded in paraffin wax extending 3 feet in front of the downstream antenna. Wax is a radio-transparent medium with index 1.5 at GHz frequencies, producing radio Cherenkov radiation with relativistic proton beams. In Set-W a 16 cm long lead target was also placed at the beginning of the tank to generate a hadronic shower. The target length was calculated to produce about one nuclear interaction on average. Yet some events will have more than one interaction, and previous studies [8]
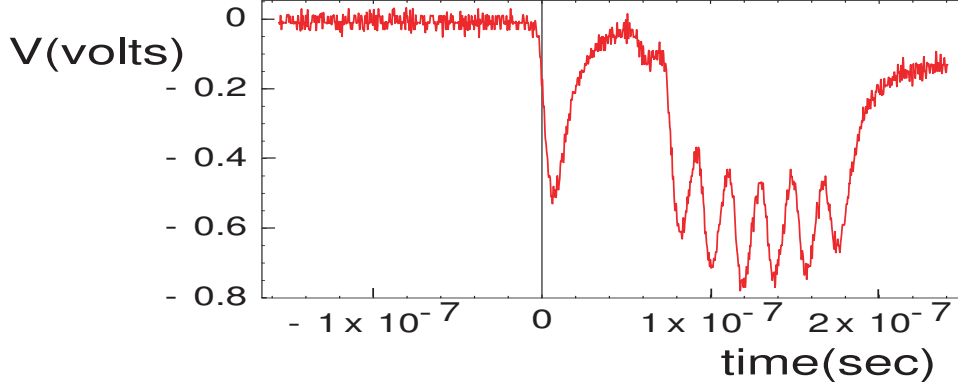
Fig. 4. The bucket structure determined from a typical phototube response. The average response over hundreds of runs is practically indistinguishable. Units on the horizontal axis are seconds, the vertical axis in volts.

have shown that radio frequency radiation from fully developed hadronic showers eventually imitate electromagnetic ones. The pre-radiator was designed to produce a modest signal enhancement of order one on average, with occasional fluctuations that we hoped might stand out when data was processed.

However the central spot of the beam traversed the tank axis with fluctuations of order 1 inch, as surveyed directly with scintillator finger counters. Beam variations themselves created an order of magnitude uncertainty in the normalization of the expected signal. A plot of typical phototube response showing the bucket structure is shown in Fig. 4. We do not use the phototube response and click runs *per se* in our data processing. They are building blocks for a simulation made separately that serves as an order of magnitude consistency check which happened to work out quite satisfactorily.

The minimum rms noise values of around 30mV, compared to signal level expectations of order 100 μV, pushed the experiment into a regime of very low signal to noise (S/N). Visual inspection of the trigger traces rejected events with grossly fluctuating or incomplete buckets. The remaining events were averaged to construct $\bar{V}(t)$ in which the rms noise was reduced to approximately 1.5mV. Fig. 5 shows the averages in the two antennas of 361 good runs in air. Even after averaging there is no visible trace of a signal beginning near the center of the time window.

1.1.3 *Observations During the Run*

Careful visual observation played a role during running. Good, well centered events were accompanied by a barely discernable 53 MHz structure we called "picket fences." Picket fences mysteriously survived explicit soft-
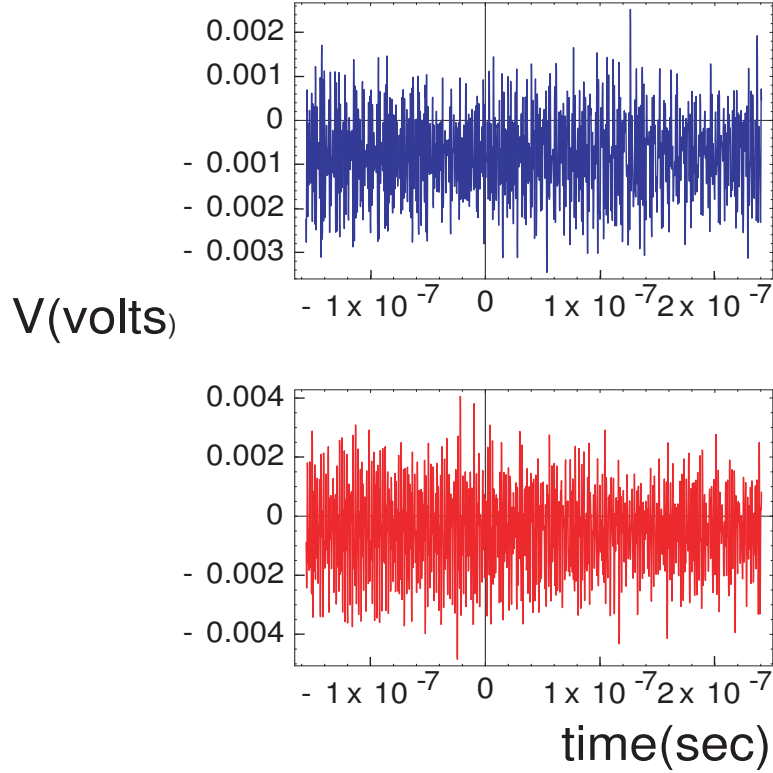
Fig. 5. The average voltages of channel 1 (top) and channel 2 (bottom) for 361 good runs in air. There is no visible trace of a signal beginning near the center of the time window. Units on the horizontal axis are seconds, the vertical axis in volts. The vertical black line shows the trigger location.

ware filtering which deleted the 53 MHz region, and were eventually identified as *beat frequencies*, due to the product of the bucket form factor's 18.9 ns repetitions and the transfer function. By picket fence observation we could center the beam to about 1 inch accuracy, an empirical detection without need for signal processing. When later classifying events using offline software, picket fence behavior was also found in 30-100 noise files, wherein pickets extended across the whole range of the data, including the pre-causal onset region. Subsequent consultation with the Fermilab staff[11] confirmed the likely existence of a precursor associated with the formation of buckets and occurring whether or not we received the beam. Removal of these files made no difference to signal detection in the causal onset region except for dilution of statistics. They also cannot be an explanation of our observations in the signal region because the pickets extended over the entire data record, noise-region and signal region alike. Beam centering done with the beam-on picket-fences actually made an unexpected consistency check. In these runs a hypothetical phototube signal would remain the same, as the trigger scintillator paddle was bigger than the beam displacement.

In Section 2 we will describe the analysis strategy we devised to extract a signal under high noise conditions. To simplify the presentation we briefly summarize some results.
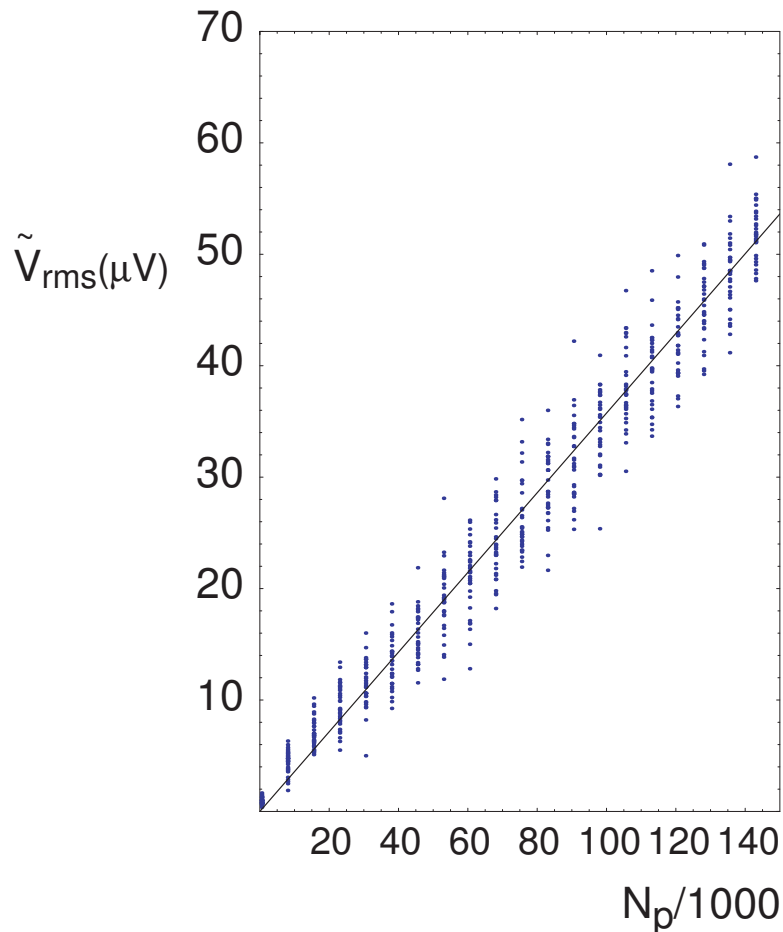

Fig. 6. Linear scaling in the signal region. Voltage versus particle number N_p, in thousands of protons, normalized to 500 protons per run. The square-root random-walk contribution has been subtracted. The slope of the straight line fit has been predicted by the simulation within a factor of 2. Air Ch 1.

Figure 6 shows the filtered *rms* voltage $\tilde{V}_{rms}$ extracted from the signal region of the data, as predicted by causal arrival of the fields moving with the beam. The figure shows $\tilde{V}_{rms}$ as a function of the number of protons N_p used in the analysis. The linear dependence of $\tilde{V}_{rms}$ on N_p is evidence of *coherence*, by which the voltage measured is proportional to the total charge producing the field. The slope of the plot is directly proportional to the number of particles per run, the amplifier gains, antenna response, and finally the elementary proton charge e.
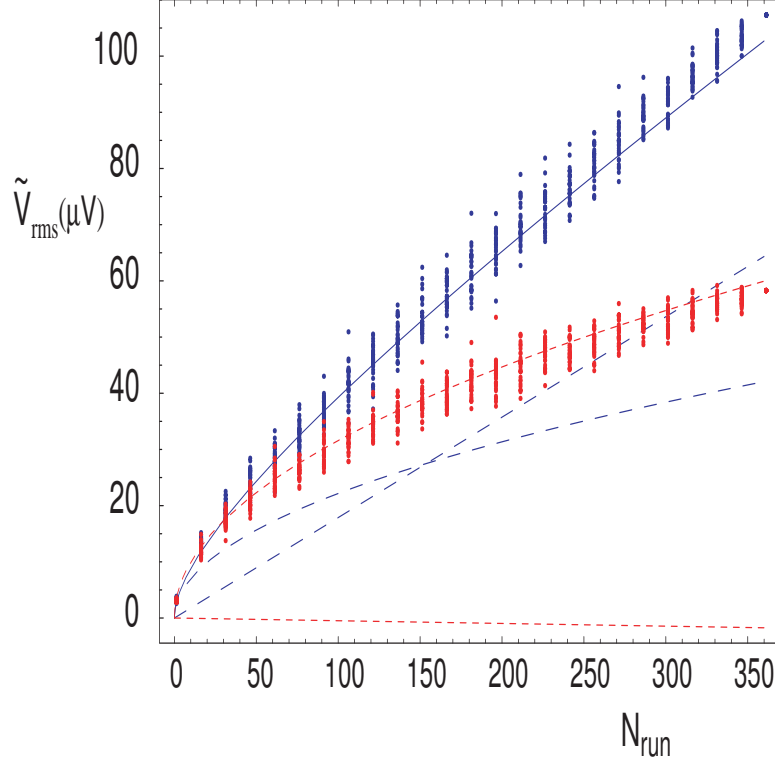
Fig. 7. Dependence of *rms* filtered voltage $\tilde{V}_{rms}$ on the number of particles, as represented by the number of randomly-permuted runs N_{runs} analyzed. Linear coherent scaling with N_{runs} of the signal region (top curve, blue online) is evident. Scaling with $N_{runs}^{1/2}$ occurs for the noise region (bottom curve, red online). The separate contributions of $\alpha N_{runs}^{1/2}$ and βN_{runs} are shown for comparison (dashed lines). Air Ch 1.

In making this figure a term scaling like $N_p^{1/2}$ and consistent with the noise contribution has been subtracted. Fig. 7 illustrates the separation into noise and signal components using $N_{runs} = N_p/500$. Noise contributions are consistent whether they are obtained from a fit to the signal region or whether obtained from the pre-onset parts of the data ("noise region"). The quantity $\tilde{V}_{rms}$ itself has been developed by signal/noise improvement procedure (Section 2.1) that sorts and then rejects dominant timing patterns found among correlations of the noise region. Finally (Section 2.3.2) the timing pattern of the signal region is consistent with the pattern predicted by the simulation, using a χ^2 test.

The overall normalization of our simulation is rather uncertain, and deliberately does not involve a very refined procedure. Factors contributing to the uncertainty are the beam-antenna separation, the number of protons per bucket, and the antenna-amplifier gains. Several factors multiply and it is difficult to control the normalization to better than a factor of 10 or so. Nevertheless we went through the exercise and committed to a value before examining the data. As Fig. 17 in Section 3.2 shows, the value of the slope

13

of Fig.6 has been predicted by the simulation up to one re-normalization factor $n \sim 2$.

All four channels show behavior very similar to the one illustrated in Fig. 6, 7 highlighting Air Ch 1. We were not surprised to see similar linear scaling in both Air and Wax, but a word of explanation might be in order. As emphasized earlier, the phenomenon of Cherekov radiation and the impulsive fields of virtual free space radiation are not fundamentally different in the near zone. Linear coherence is expected in all cases because the total fields are proportional to the total charge. Only at distances of "far zone" and beyond does the small virtuality of fields in Air lead to their confinement, or to propagation out to infinity in Wax.

It is interesting that by extrapolation to one particle the experiment can then measure the charge of single relativistic protons. The extrapolation may seem bold, bit it is justified once we have established linear behavior on the number of particles. Our experiment is like Millikan's with larger relative errors. The data analysis reached its limit when the statistical errors are smaller than the irreducible systematic errors. Using the normalization of the simulation predicted, and adding a generous error of a factor of 2 on the predicted slope, the experiment yields an overall measurement of the single proton's charge of order $(2 \pm 8) \times 10^{-19}$ C.

2 Analysis

Given low S/N it was necessary to devise new strategies to resolve a signal. In the following we provide the conceptual background. This will be followed by description in depth of a novel data filtering strategy:

- We sought a method not intrinsically tied to Monte Carlo simulations. The role of the Monte Carlo is to provide a secondary consistency check, and not to be the primary basis of declaring a signal.
- We realized that extensive "noise files" taken during the run should be used to define a signal self-consistently and in a mode of data-versus-data, not data versus theory.
- We sought a method that would not be overly sensitive to small variations of cuts, parameter choices, or subjective judgments, while still creating latitude for judgment to be used.
- We desired a method based on a "blind" procedure, treating signal data and noise under one uniform method without using the signal data to define a signal.

The criteria are carried out by methods based on "filters," which broadly mean algorithms discriminating on the basis of patterns in the data sequence.

A huge literature exists on "matched filters." Matched filters preferentially pass patterns of a predetermined class. Analog radio receivers essentially operate with Fourier-component matched filters. They are successful under very low S/N when signal lies in narrow frequency components, to which the radio oscillator responds resonantly.

Yet under conditions of low signal to noise, the entire strategy of "pattern acceptance veto" is unstable. By definition random noise generates all possible patterns. When noise fluctuations that happen to reproduce selected pattern overwhelm the signal, the strategy will fail. For our experiment the impulsive signal has a broad Fourier spectrum, and is also difficult to model with high accuracy. The uncertainties from beam-location, bucket fluctuations, and system calibration make a Monte Carlo-based matched filter "tuned to signal" unwise.

We found a new strategy realizing that our noise (and the noise of most experiments) is far from thermal and contains myriad correlations. Noise correlations can be put to advantage. We devised a method to systematically create "noise filters " well-matched to the patterns in noise. This requires more than one filter, but by exploiting linear combinations of patterns, even a few noise filters can capture an infinite number of noise configurations. After classifying the noise we created a noise rejection veto 100 % efficient against the patterns found. No attempt was made to select a signal pattern: instead, those components untypical of noise were simply left untouched. While some signal will be rejected, the signal/noise ratio can be substantially improved by rejecting relatively more noise than signal.

2.2 Constructing Filters Using Noise Against Noise

First define a few terms:

- A one-dimensional data list of length D is considered a vector on D dimensions.
- A *filter* is an operation on the list to return a new list. We only consider linear operations. Any linear operation on the list can be represented as multiplication by a $D \times D$ matrix.

- A "pattern" is a vector orthogonal to other patterns, and thereby independent. Orthogonality is defined using the dot product $\langle A \mid B \rangle = \sum_j A_j B_j$.
- On D dimensions there are at most D patterns, which can be used to make an orthonormal basis.
- The set of all linear combinations of a subset of patterns, or basis vectors, is called a subspace.
- Filters can accept, reject, or manipulate patterns and thereby control subspaces in which data occurs.

2.2.1 Extracting Pattern Subspaces

Consider summing many instances (label J) of a data vector d_i^J that is "trying to repeat" a single pattern p_i. If patterns occur with random coefficients α^J, then

$$d_i^J = \alpha_J p_i,$$

and the sum tends to zero. If the sum happens to be non-zero it can be recorded. This still leaves all the patterns that cancel undetermined.

The solution to discovering general data patterns is to make *outer* products, namely matrices, and average them. Take a single data string d_i, and make the matrix $d_i d_j$. Suppose another data string happens to be $-d_i$; it yields the same matrix $d_i d_j$. Regardless of coefficients similar patterns *add* upon averaging the outer products.

Summing outer products produces a density matrix[2] ρ_{ij},

$$\rho_{ij} = \sum_J \alpha_J^2 p_i p_j \to constant \times p_i p_j.$$

The underlying pattern p_i is then recovered as the eigenvector of ρ_{ij}.

When more patterns with arbitrary coefficients exist, a set of patterns that make an optimal basis is again found from the eigenvectors of ρ. To make an "optimal" filter, then: (1) develop the optimal patterns from the density matrix of a sufficiently large sample; (2) classify pattern importance quantitatively; (3) expand the data in the basis of patterns (4) throw away components not desired (normally small, or rare elements); (5) revert the expansion back into the original components.

---

[2] The term borrowed from quantum mechanics is exact. A study of quantum theory led to the procedure.

2.2.1.1 White Noise: Textbook "white noise", often called simply "noise", is defined by correlations of data n_i that has a density matrix

$$< n_i n_j >= \sigma^2 \delta_{ij}.$$

The statistical average $< >$ is probed experimentally by adding up noise samples n_i^J. If physical noise would actually satisfy the textbook criteria, then no particular pattern is favored. All possible patterns are produced equally. The attempt to extract a special eigenvector of δ_{ij} yields no particular vector because all vectors are eigenvectors of the unit operator.

Ideal white noise is actually rare. For this reason one can classify correlations in the noise to find the dominant patterns that actually exist. When $S/N >> 1$ patterns in the data will overlap with signal. Then retaining the dominant patterns improves S/N. We faced a situation of $S/N << 1$ that naturally suggested the reverse procedure, and also necessitated optimizing our filters.

2.2.2 Mathematical Description

Our filters use a variation of the Karhunen-Loeve(*KL*)[9,10] method from image processing. Our reverse *KLJ* filtering is done in three steps: (1) Data from the noise region of the files (the first 900 points) are partitioned into non-overlapping segments of length D ("bins") containing vectors $\left| noise^J \right\rangle$. (2) We construct orthogonal projectors π^a such that

$$\frac{< \langle noise | \pi^a | noise \rangle >}{< \langle noise | noise \rangle >} = n(a) \rightarrow max,$$

where $< >$ denotes the sum over bins. Each projector π^a defines a one dimensional "noise subspace."

Optimization is done by solving an eigenvalue equation[9]. Construct the noise density matrix

$$\rho_{noise} = \sum_J \left| noise^J \right\rangle \left\langle noise^J \right|.$$

Solve the eigenvalue equation

$$\rho_{noise} \left| e^a \right\rangle = n^a \left| e^a \right\rangle.$$

By the Rayleigh-Ritz variational theorem, the eigenvectors maximize $\langle e| \rho_{noise} |e\rangle$, for any possible normalized $|e\rangle$. We will call the normalized eigenvectors "noise states" for simplicity. Assigning $\pi^a = |e^a\rangle \langle e^a|$ then gives

$$\langle e| \rho_{noise} |e\rangle = tr(\pi^a \rho_{noise}) = n^a,$$

where tr is the trace. The optimal noise subspaces so constructed have the maximum overlap with noise state-by-state.

The idea of optimal filtering itself is not new. The process of *Wiener filtering* is invariably described as optimal. It is easy to show that when noise is "ideal", featureless, and time-translationally invariant, then the noise states will be Fourier modes. For our purposes it is unwise, inefficient and unnecessary to make such idealized assumptions when one has actual data for the noise.

We label the noise states by their eigenvalues, sorted from largest noise power to smallest, $n(1) > n(2) > ...n(D)$. We then construct $\pi_{noise}(\mathcal{R}) = \sum_a^{\mathcal{R}} \pi^a$, which is the most efficient noise-passing filter of given rank $\mathcal{R} \leq D$. (3) Our filter consists of applying $\pi_{signal}(\mathcal{R}) = 1 - \pi_{noise}(\mathcal{R})$ to the data, removing $\mathcal{R}$ dimensions of noise and passing $D - \mathcal{R}$ dimensions populated by noise as little as possible. In symbols,

$$|V\rangle_{filtered} = (1 - \sum_a |e^a\rangle \langle e^a|) |V\rangle.$$

We scrupulously arranged that the filter construction never uses the signal region of the data. Under the hypothesis that the entire data set is noise, the filter will not favor the causal onset region over any other.

2.2.3   Indices and Translational Properties

Partitioning the data into "bins" of length D loses no information, and is simply a relabeling of indices:

$$data_i \rightarrow data_k^J; \quad J = int(\frac{i}{D}); \quad k = mod_D(i).$$

Here *int* takes the integer part of its argument, producing the bin-index J. The function mod_D is the remainder of division by D, yielding the index within the bin k. The inverse of the transformation assigns index i by the rule

$$i = JD + k.$$

In this way filtered bins are conjoined to re-make data in its original index notation.

Generally the noise states make a complete set for any data in each bin or length D. [3] Thus repeated application of the complete noise state basis bin-by-bin is perfectly lossless. When a filter is applied one deliberately rejects information *within bins*, without changing the bin dimension D. To make this less abstract, suppose the filter only retains 2 specific Fourier modes on a $D = 10$ dimensional bin. Suppose the data length is 2000 points. Then on each of 200 bins, those components of the data in each Fourier mode will pass. Experience processing physical data shows that one overall mean - the zero frequency mode - should be removed, else it overwhelm all other modes.

After filtering the labels J, k can be used, or they can be reverted to the original monotonically running i:

$$data_i^{filtered} = \sum_{Jk} \delta(i - J D + k) \, data_k^{J, \, filtered}. \tag{4}$$

At this point the partitioning and binning has done its job and "disappeared" leaving revised data in its original format.

A question arises whether the filtering process depends on the start-point of the bins. The answer depends on the nature of the data used to make the filter, and the number of states retained. The extreme case of retaining one lone state, for example, forces one specific pattern to re-appear in each bin with sign and normalization that is the best possible fit by a very limited subspace. Retaining many states has a strong tendency to be nearly translationally invariant and independent of how binning is started. One explanation is that special glitches in data between bins, that might distinguish one bin-start point compared to another, tend to be small effects compared to the accumulation of data over the domain of many bins. The other explanation seems to be that translational symmetry of data correlations is a generically reasonable approximation. Even so, forcing Fourier modes (Wiener filtering) omits phase correlations that cannot be seen in Fourier power spectra and will degrade performance.

---

[3] An exception occurs if fewer vectors than the length D are used to make noise states, or if those vectors for some reason lack linearly independent components. This is readily cured by adding more vectors from the noise region.

2.2.4 Timing Resolution

Noise states tend to be ordered in "smoothness" and are often similar to Fourier modes. Filtering removes components needed to make all possible patterns. Filtering is then guaranteed to downgrade timing resolution within bins. The proof of this is very simple. Consider the basis with elements that are spikes at each point k: $e_k^a \to \delta_{ak}$. Aligning a data element perfectly with such a spike gives perfect timing resolution, namely the sampling time Δt. A superposition of many noise-basis elements is generally needed to make spikes, as in Fourier series. Decimating the basis by filter construction can only downgrade timing resolution compared to examining data spike-by-spike. Conversely, triggering spike-by-spike on raw data allows the maximum bandwidth of noise to intrude, and will generally produce the lowest possible S/N. Evidently an "uncertainty principle" operates which is more general than the usual relationship for the Fourier domain.

Given the large amount of noise our filters reject, timing resolution inside bins is quite poor, and timing resolution of order $D\Delta t$ is expected. We made several studies hoping to evade this reality before we realized that the extended time structure of the buckets made it pointless. We turn to more discussion on how filter parameters were chosen.

2.2.5 Signal Retention

Our method can increase signal to noise when sufficient signal exists in the subspace retained. One also wants to make sure that tentative signals lifted out of the noise are not overly sensitive to fine details of procedure. These facts determine the useful ranges of bin dimension D and the filter rank $\mathcal{R}$.

We studied simulated signal events under the actual filters. Our nominal signal is $V_\omega^{signal} = nN_p T_\omega^{click} \bar{F}_\omega$, where n is a normalization adjustment relative to theory, and $\bar{F}_\omega$ is the Fourier transform of the averaged phototube pulse. In the time domain $\bar{F}(t)$ is a long bump with one initial and six subsequent bucket peaks repeating at 18.9 ns. This form factor and the need to average data with timing jitter precluded sharp timing resolution that would simplify analysis. The bin length D was chosen to compromise between noise rejection (improving with larger D) and timing resolution (deteriorating with larger D). For $D \lesssim 25$ our signal simulation showed predominant overlap between noise and signal and the filter is too weak. For $D \gtrsim 28$ separation is good. Otherwise dependence on D is weak, and we settled on $D = 32$. The rank $\mathcal{R}$ is determined from the signal efficiency $\eta(dB) = 10\log_{10}(\sigma_{passed}/\sigma_{raw})$, where σ denotes the *rms* of noise regions passed by the filter compared to raw data. Set by set, we adjust $\mathcal{R}$ to be as
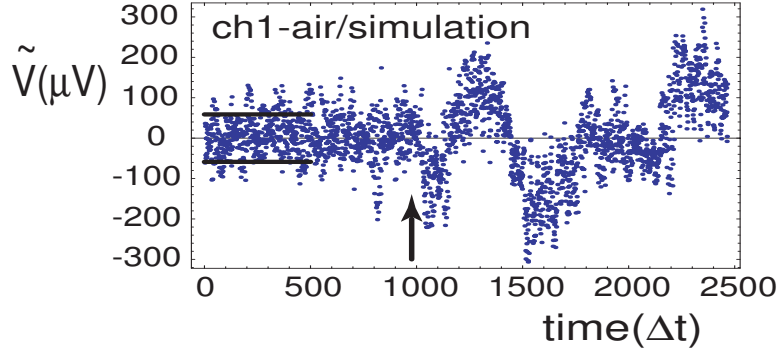
Fig. 8. Simulation of signal shape post-filtering, channel 1(air). Earliest signal onset occurs at the arrow (point 970); shape is dominated by the form factor. Units are μV; time in units of Δt=0.16ns. Horizontal lines show $\pm 1\sigma$ in the noise region.

large as possible to reject noise. Meanwhile $\eta(dB)$ is required to be rather flat, so that signal efficiency is not overly sensitive to the exact choice of $\mathcal{R}$. We insist $\mathcal{R} \rightarrow \mathcal{R} \pm 1$ does not change $\eta(dB)$ by more than about 1 unit. This procedure fixed $\mathcal{R} = 11$ for Channel 2 in Set-A and $\mathcal{R} = 12$ otherwise. A simulated signal ($n = 3$) plus noise $\tilde{V}^{signal}$ post-filtering is shown in Fig. 8. Noise was obtained from pre-onset regions of the data to make this figure. The filter retains over 95% of the signal power while reducing the noise by a factor of about 20. None of this is very sensitive to the details and uncertainty in modeling the signal. It is a generic fact that almost any signal has components different from the details of the physical noise. To achieve similar results without the noise-rejection filter would require about 400 times more data, the equivalent of running for about 3 years.

Dependence of the filtering process on noise states retained is shown in Fig. 9. The left column shows data processed by filters that keep noise state labels equal to or exceeding a cutoff N_{cut}, shown at top of frames. Thus $N_{cut} = 6$, used for the middle panels in both columns, use filters which omit the first 5 most important noise states. The right column shows the same procedure applied to the form factor, as calculated from click file and average bucket structure. Since the means are removed before filtering, as discussed earlier, one constant voltage offset shifts the figures. Note changes in the vertical scale of the data, especially compared to the simulated signal. Preferential passage of the signal compared to noise is clear, increasing the S/N ratio.

A rather different illustration is given by Fig. 10. The left panels show the

action on the physical noise on a typical file. The right panels shows the effect on simulated (textbook) white noise from a Gaussian distribution with the same standard deviation. Filtering of the physical noise is efficient, while extremely mild on Gaussian random numbers. As explained earlier *perfectly random* numbers fill out all possible patterns and tend to pass any conceivable filter. Meanwhile *physically random* data have correlations that allows far more effective rejection.
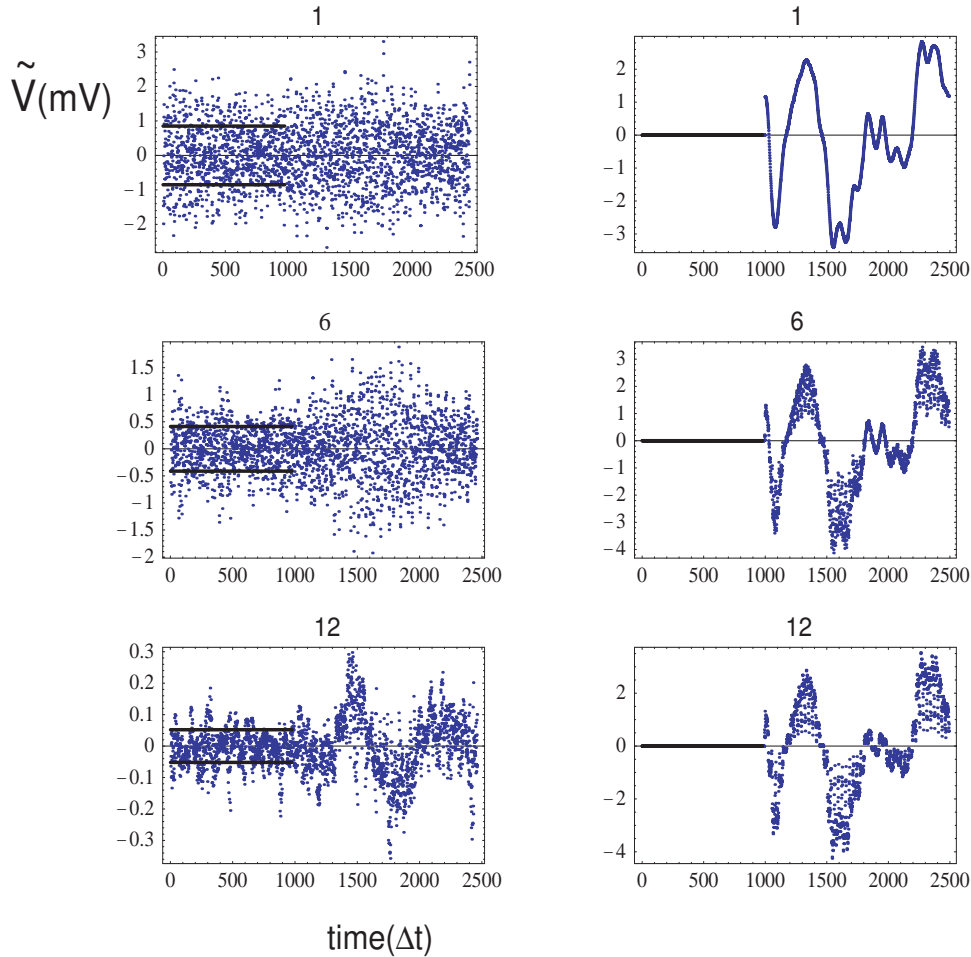

Fig. 9. Effects of filter subspaces. Filters retain noise-state labels equal to or exceeding N_{cut}, shown at top of frames. The top rows labeled "1" keep all states, the bottom keep $N > 12$. *Left column*: Filtered data (air). Note changes in vertical scale from different filter states. *Right column*: The filtered form factor, as calculated from a click file and average bucket structure, and using same procedure as the data. Preferential passage of the signal compared to noise is clear. Horizontal lines consistently show $\pm 1\sigma$ computed from the first 950 points after filtering. A constant voltage is removed in filtering, and has not been restored.
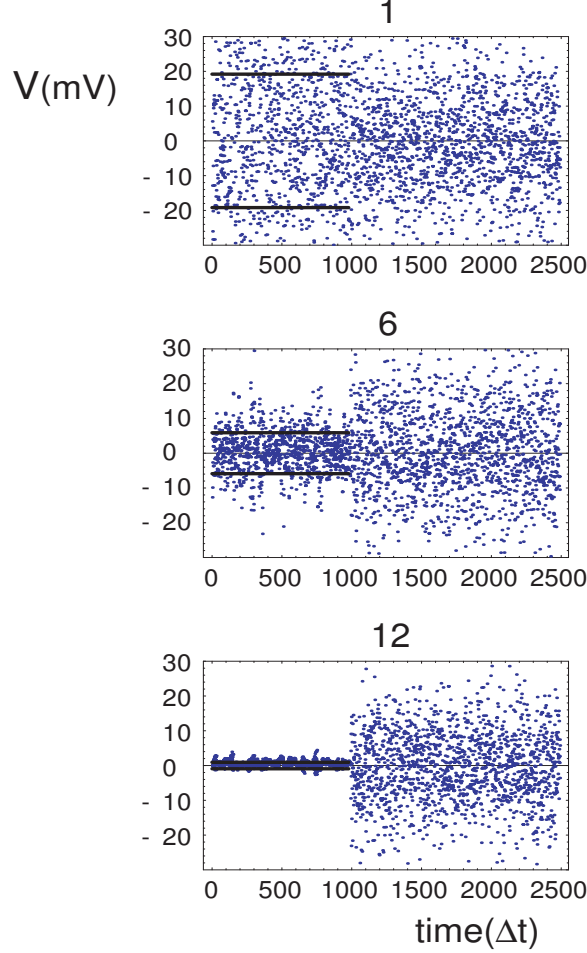
Fig. 10. Effects of the filter, as made with different subspaces, acting on the physical data compared to randomly generated numbers. Labels as in Fig. 9. The first 950 points come from a typical data file; the last 1250 points come from a Gaussian distribution (idealized noise) with the same standard deviation. Physical noise is rejected far more effectively than idealized noise. Horizontal lines show $\pm 1\,\sigma$ over the regions indicated.

2.2.6 Simultaneous Mode and Time Information

Our filtering scheme has a rather natural graphical representation. For each given bin-time, projections of the data into each noise mode ($data_k^{J,\,filtered}$, Eq. 4) are plotted on the vertical axis of "running mode" plots (Fig. 11). The noise subspaces are ordered from bottom (maximum noise) to top (least noise). This is repeated for each bin time plotted along the horizontal axis. A graphics program then constructs contours of constant amplitude in the mode-versus-time plane. Rejection of a noise subspace consists of ignoring data below a horizontal line (subspace π_{noise}). Detection of a signal within a bin-time or two consists of observing structure in the region above the line (subspace π_{signal}.)
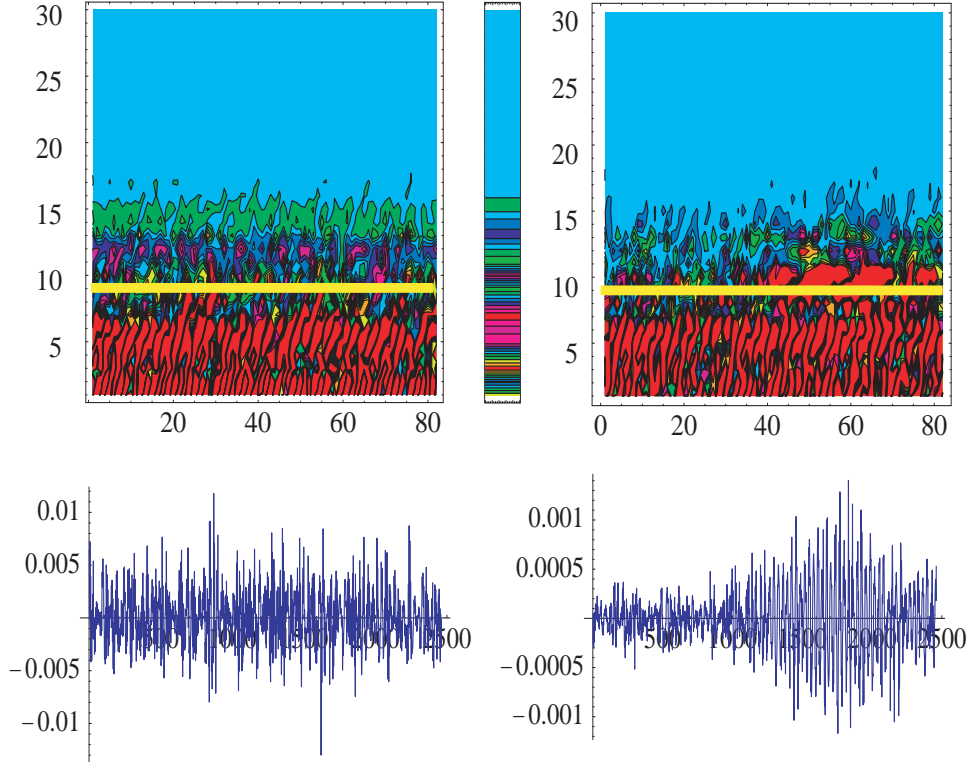
23

Fig. 11. Simultaneous mode and time information. Upper panels show noise-basis projections of the data arranged by rows, from bottom (most noise) to top (least noise). Binned-time is plotted along the horizontal. Visual appearance suggests keeping components above the horizontal line for R> 9, rejecting integrated power at the 200:1 level, which has been done here to demonstrate insensitivity to fine details of the cut. The thin middle panel shows the single-bin signal click-pattern in the noise basis, expanded horizontally to be visible. Bottom panels show the data of a filtered typical run (left) and the filtered average over runs (right), ch 1.

We also included a thin middle panel (Fig. 11) showing a single-bin signal click-pattern in the noise basis. It has been graphically expanded horizontally to be visible. The panel demonstrates that signal protrudes well out of the noise subspace cut by the horizontal line.

The running mode plot then shows all the data, for all times, organized into noisy and quiet projections. It is a matter of taste whether or not to square the mode-projections to examine mode power, to plot positive and negative values scaled to make structure visible, to take the logarithm, and so on. Many options exist. Fig. 11 also makes visible the effects of adjusting the "cut" on noise power. Visual inspection suggests making a cut with R> 9, weaker than the ones we chose on a quantitative basis. (This sort of visual inspection is *retrospective* towards finding a signal and might have to be justified.) The figure shows action of the weaker filter is very similar to our analysis, because we did not choose our cuts in an overly sensitive region.
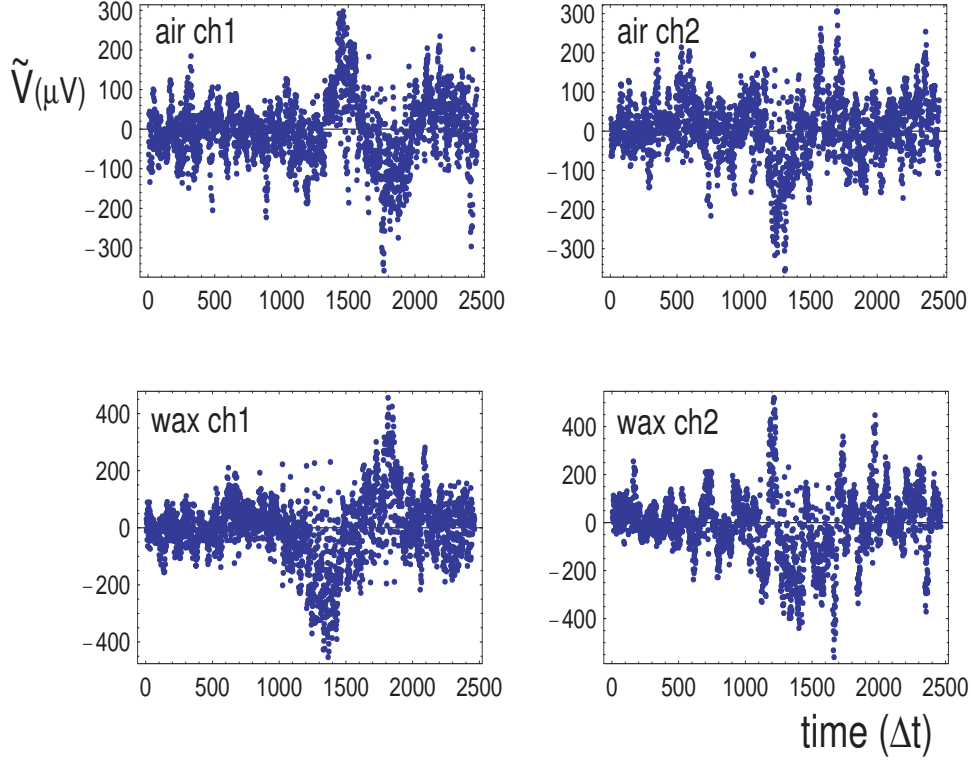
Fig. 12. Output traces for filtered data sets $\tilde{V}$. Free space medium ("air", top) and wax (bottom), for channel 1 (upstream, left panels) and channel 2 (downstream, right panels). Causal onset can be seen in the signal region after the center of each dataset, as in Fig. 8. Voltages in μV; time in units of Δt=0.16ns.

2.3 Basic Analysis of Data

We examined the output $\tilde{V}(t)$ of averaged data for each dataset after filtering. Fig. 12 shows the filtered traces in two channels for the "air" and "wax" cases. Recall that the air data is simply a beam traveling through empty space, and the wax data includes a lead pre-radiator that generates a hadronic shower. Visual inspection shows a difference between the initial region and the region associated with causal arrival of the beam. (The earliest causal onset point, as determined from our phototube and click-file calibration, is point 970 for channel 1.) Since the signal to noise was very poor to begin with, we contented ourselves with simple statistical methods to quantify detection.

2.3.1 Data Distributions

We divided the filtered data into two regimes, "before" the causal onset point (the first 970 points), and "after" (1490 points). The total number of 2460 points comes from the filter retaining 77 bins of 32 points/bin. The

channel	σ_{before} (N_{before})	σ_{after} (N_{after})	$N(\tilde{V}_{before} > 3\sigma_{before})$	$N(\tilde{V}_{after} > 3\sigma_{before})$
Air Ch 1	59 (970)	105 (1460)	10	166
Air Ch 2	74 (970)	103 (1460)	16	91
Wax Ch 1	67 (970)	150 (1460)	8	284
Wax Ch 2	81 (970)	170 (1460)	6	266

Table 1
Basic statistics of the filtered data signal $\tilde{V}$ before causal onset and after. Standard deviations (σ, in units of μV), number of points (N), and number of points exceeding 3σ in the four channels measured.

standard deviation σ of the data sets before and after onset are recorded in Table 1. Several sets show hundreds of points in the post-onset region with amplitudes larger than $3\sigma_{before}$ (Table 1). These statistics are evidence for events related to the arrival of the beam in both the air (virtual radiation) and wax (real radiation) cases.

More information is given in Fig. 13. The figure compares histograms of the filtered data in the before and after regions. The distributions are reasonably consistent with Gaussian forms, but there is insufficient statistics to address the crucial issue of behavior in the tails. The histograms support the information given by the standard deviations, namely that the distributions after onset are wider than before. We estimated P-values for the data before onset to fluctuate to the degree seen after onset. If one assumes Gaussian distributions the P-values for all four cases were less than 10^{-9}. P-values depend on the number of degrees of freedom, for which we used $20/32$ of the number of points to account for the reduction in freedoms from filtering.

2.3.2 Time Structure

The statistics just cited are "bulk" integrated measures that contain no information on the time-structure of the data. We attempted to quantify agreement of the shape of the filtered data with the filtered simulation, as follows.

Recall that $\tilde{V}_t^{signal}$ is the filtered simulation, and $\tilde{V}_t$ is the filtered data. Remove the means and normalize each vector. Then $C(0) = \sum_t \tilde{V}_t \tilde{V}_t^{signal}$ represents the dot product of the two vectors, which is a measure of agreement of the shapes. Perfect agreement corresponds to $C(0) = 1$, which is statistically very unlikely for vectors with many components. More generally, one can shift one pattern relative to another by defining running correla-
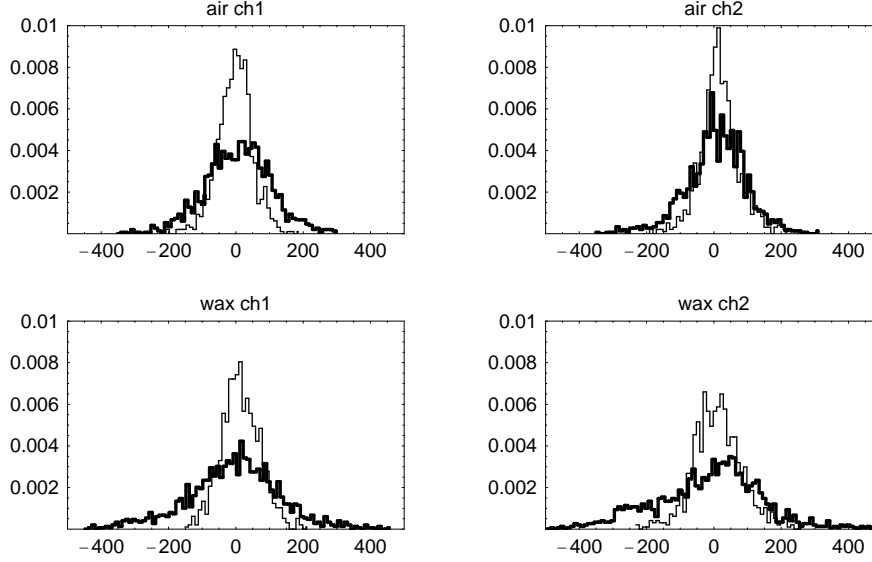
Fig. 13. Histograms of filtered data $\tilde{V}$ in μV from the time segments *before* causal onset (light lines) and *after* causal onset (heavy lines), see text. The four panels are for the cases of free space medium ("air", top) and wax (bottom), for channel 1 (upstream, left panels) and channel 2 (downstream, right panels).

tions $C(T) = \sum_t \tilde{V}_{t+T} \tilde{V}_t^{signal}$, which happens to coincide with the 2-point correlation of classical statistics.

The width of maximum $C(T)$ depends on the detailed way in which it is calculated. For one thing, a very extended pattern degrades timing information, while tending to increase statistical significance. Generally our studies yielded a timing resolution of order 200 sample time units (32 ns), which was consistent with our simulation. The timing resolution after filtering is too poor to resolve the time separation between the two antennas to provide evidence of causality. However the observation in the causal onset region can only come due to fields traveling with the beam. Observation of 2 traces with consistent timing, shape, and size predicted by simulations provides evidence that the experiment detected virtual radiation.

We constructed our simulation in the simplest way possible because we have minimal information on the fluctuations of the beam position event-to-event, and also the bucket structure occasionally changed significantly. Rather than throw away data it was to our advantage to make averages. When we made running correlations $C(T)$ of our filtered data with the simulations, we found evidence supporting good agreement of the simulation time structure and data time structure in the form of large correlations $C(T) \gtrsim 0.4 - 0.5$ for simulation lengths of 800 points. This was seen in each of the 4 data sets, and similar results were seen for many different simulation lengths. The naive statistical probability one might estimate for

27

the results in a random sample were vanishingly small in many cases. At the same time, we found non-negligible correlations of some data sets in regions of T that were supposed to correspond to noise. This is explained in part by the fact that the filter introduces correlations by retaining a class of data patterns that is not ideally random. The running correlations studies also depend on the length of the simulation region, the bin dimension and rank D, R, and the length of offset T, making for a very complicated statistics problem. For this reason, we did not pursue a full statistical analysis of $C(T)$ further.

This analysis gives evidence for detection based on appearance of signals in the causal region. Noise fluctuations have been ruled out. There remains only a possible RF background, which if postulated must be closely timed event by event, and also pass the filter. The Fourier spectrum of the phototube pulse was measured and is predominantly below 100 MHz, explaining our 200 MHz high-pass filter choice. Lack of complete immunity to phototube noise was clear in "whopper" events when the beam was steered inadvertently into the phototube itself. These events were thrown out early. However, we also quantified phototube noise with runs with finger-counter triggers to eliminate this possibility of our trigger signal polluting our antenna data.

To quantify the signal further we turn to evidence of coherence of the signal.

3   Coherent Scaling

The hallmark of coherent Cherenkov radiation is linear scaling of the signal with the total number of particles. To test linearity one might vary the total number of charged particles in the beam. However it is unrealistic to ask the test beam facility to adjust particle number while maintaining a consistent time-structure. We varied the particle number by adjusting the number of runs included in the data analysis.

We sum the filtered voltages for a given number of runs N_{runs}. We then calculate the *rms* of the result, producing

$$\tilde{V}_{rms}(N_{runs}) = rms\left(\sum_{k}^{N_{runs}} \tilde{V}_{k}, \right),$$

where k is the run number and symbol *rms* takes the standard deviation. The results of adding a subset of runs depends on the order in which runs are selected. To make an unbiased sample we repeated everything 30 times

with the order of runs randomly permuted. Fig. 14 shows the results using the *signal region* (the last 1500 points) for the four cases of channel 1, 2 in Air and Wax. The figures show evidence of linear scaling, *i. e.* coherence, both in the Wax and Air cases.
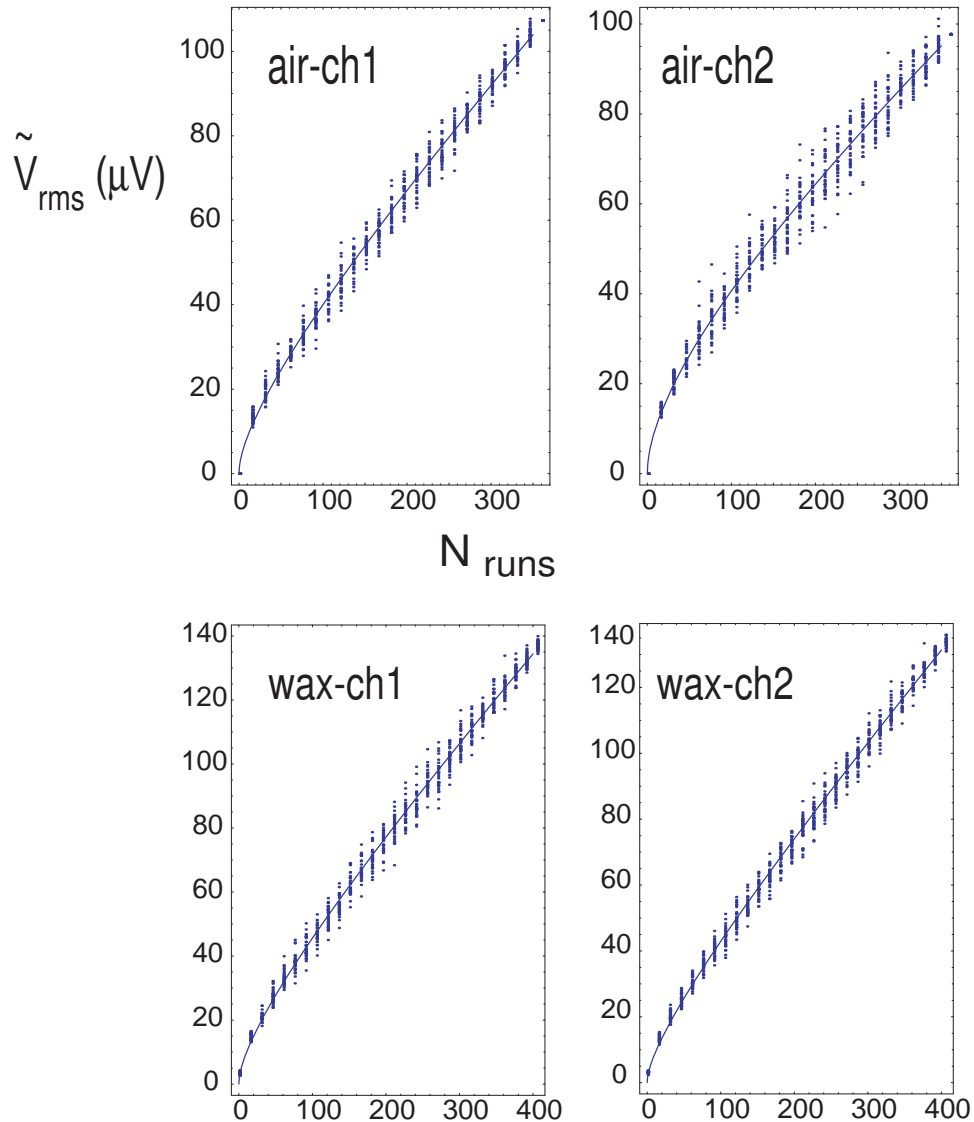

Fig. 14. Dependence of *rms* filtered voltage $\tilde{V}_{rms}$ on the number of particles, as represented by the number of randomly-permuted runs N_{runs} analyzed. Data from the *signal* region, as defined in the text. Curves are $\tilde{V}_{rms}(N_{runs}) = \alpha N^{1/2}$ and βN, with parameters given in Table 2.

We fit dependence of the $\tilde{V}_{rms}(N_{runs})$ on N_{runs} to an ansatz

$$\tilde{V}_{rms}(N_{runs}) = \alpha N_{runs}^{1/2} + \beta N_{runs}. \tag{5}$$

We call parameter α the *random walk* component, representing scaling with

channel	α	β
Air Ch 1	2.2	0.18
Air Ch 2	2.9	0.12
Wax Ch 1	2.4	0.22
Wax Ch 2	2.0	0.23

Table 2

Parameters of the fit $V_{rms}(N_{runs}) = \alpha N_{runs}^{1/2} + \beta N_{runs}$ developed in the signal region, as defined in the text. Units of α and β are μV. Correlated errors are discussed in the text.

$N_{runs}^{1/2}$ typical of noise. One may interpret α, which has units of μV, as a filtered noise-standard deviation, subject to fitting uncertainties. Parameter β is the *linear coherence* component. This parameter can be interpreted as the filtered electric field per 500-proton run, measured in μV via pre-factors that include antenna response and amplifier gains. For reference, our filtered simulation predicted a value of $\beta \sim 0.1 \mu V$/run, with large systematic uncertainties already cited.

Basic expectations for the α and β parameters are developed from Table 1. Air Ch 1, for example, has *rms* voltages of $\sigma_{before} = 59\,\mu$V in the region before causal onset, and $\sigma_{after} = 105\mu$V after causal onset. These figures represent the cumulative outcome of averaging 361 files. Assuming σ_{before} is nothing but random walk noise predicts $\alpha_{before} \sim 59\,\mu V/\sqrt{361} \sim 3.1\mu$V. Naively subtracting this from the signal region after onset, and assuming the balance is due to linear coherence, predicts $\beta_{after} \sim 0.15\mu$ V.

Quantitative fits to the N_{run} dependence are shown in Fig. 14. The first and last cases of N_{runs} have trivial fluctuations and were dropped, a small effect. Table 2 shows the linear coherence and random walk components are very similar for the four data sets. The fits are quite consistent with the simple argument given above. Fitting two similar powers such as $\alpha N^{1/2} + \beta N$ is known to be "ill-conditioned." Each parameter can simulate the effects of the other over a finite N range. For this reason the fit parameters need to be assessed with correlated errors. Discussion of the error ellipses is given in Section 3.1.

Thus the evidence for signal observation for the *averaged* data sets cited in Section 2.3.1 is consistent with *varying* the number of protons in the beam, extracting the term linear in the number of protons, and then averaging.

We now describe a more detailed study that develops an error estimate of parameters α, β, which will be used in calculating a χ^2 goodness-of-fit measure, described below.

Recall that Fig. 7 cited earlier shows fits for $\tilde{V}_{rms} = \alpha N_{runs}^{1/2} + \beta N_{runs}$. The top data set and curves (blue online) comes from the signal region, while the bottom data and curves (red online) come from the noise region. The random walk component is the only significant term in the noise region. This is consistent with the observed difference of noise and signal region established in Section 2.3.1. It is also consistent with the random walk contribution of the signal region, Table 2. A very small, negative (unphysical) linear term is generated by the best-fit procedure.

Each data set has comparable correlated errors. We present the details of analysis of Air Ch 1. Fig. 15 illustrates contours of goodness of fit versus parameters α, β. Goodness of fit for this figure is defined by the usual χ^2 formula, using the average of the fluctuations seen in Fig. 7 as the statistical error. Dependence on N_{runs} of $\tilde{V}_{rms}$ is separated into using the *signal region* (the last 1500 points) and the *noise region* (the first 900 points).

In principle Wax Ch 1 and Air Ch 1 should differ primarily due to the lead pre-radiator and radio cavity mode re-configuation from adding the dielectric. If hadronic fluctuations were a substantial effect we might see larger signals in the Wax case for both Ch 1 and Ch 2. The nominal Cherenkov channel is Wax Ch 2, which in theory should have smaller signals due to the factor of $1/\epsilon$ in Eq. 2. We observe that Air Ch 2 has a somewhat lower β parameter than the others. It is interesting that the linear coefficients β in Table 2) for Wax (Cherenkov radiation) are somewhat larger than those in Air (virtual radiation). We were gratified that the different cases are so comparable. First, radio-frequency normalizations are very difficult to control to the level of a few dB. Second, the response of the tank at the location of the antennas depends on the presence of the dielectric. The antenna for Ch 2 was also re-positioned and re-connected between runs. The uncertain accumulation of physical effects degrades our ability to accurately verify normalizations. For these reasons, and the ill-conditioning of power-law fits, we find the consistency of parameters seen in Table 2 quite acceptable.

The random walk/linear coherence parameter regions of the signal and noise region are well separated. As expected each fit has a substantial degeneracy in α, β parameters. Good fits to the signal region are obtained over a line $\alpha \sim 5.05 - 15.7\beta$, for $0.08 < \beta < 0.22$. Despite the degeneracy of the fit, there is no overlap of the error ellipse of the noise region with the signal

region in Fig. 15. This demonstrates that the signal region can only be fit with a dominant linear coherence behavior.
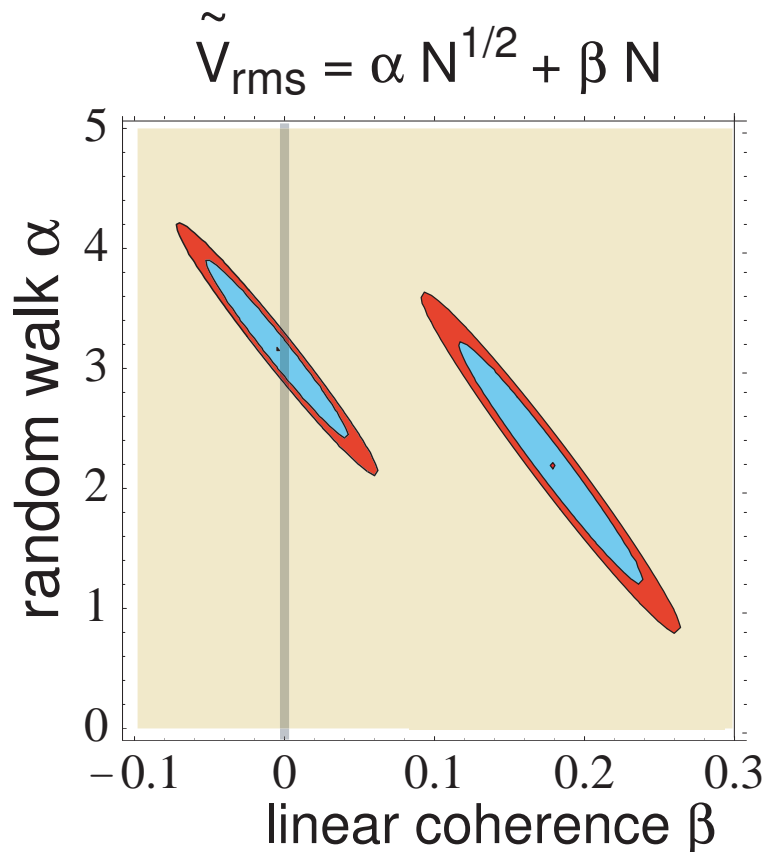

$$\tilde{V}_{rms} = \alpha \, N^{1/2} + \beta \, N$$

Fig. 15. Contour plot of χ^2 of the signal region and noise region versus parameters α, β. Contours represent unit intervals of χ^2. The minimum χ^2/dof (central dots) is close to one.

3.2 Signal Shape

Finally we quantified the shape of the filtered data set (average of all runs) compared to the filtered simulation. We gave the simulation 2 parameters, consisting of a normalization parameter n and timing offset t_0:

$$\tilde{V}_{fit}(t) = n\tilde{V}_{sim}(t - t_0).$$

The parameter n accounts for uncertainties in the normalization of the simulation and the number of particles in the beam spill. The timing offset t_0 accounts for uncertainties in timing delays, the bucket form factor, and in extra timing offset caused by the binning procedure, which introduced the discrete time bins used to make the filters.
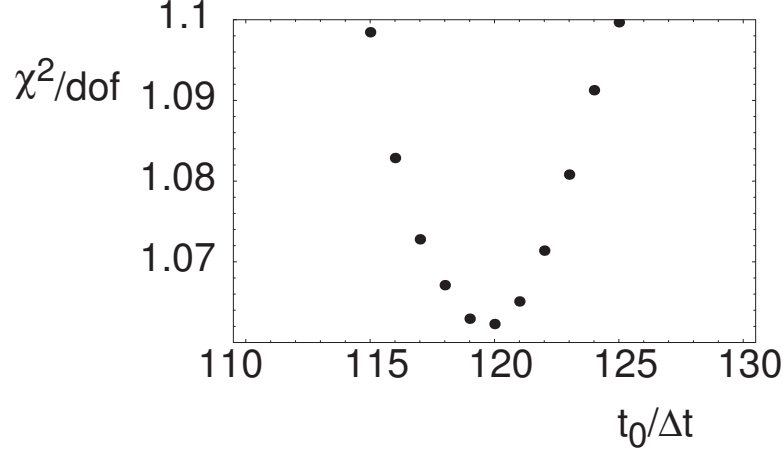
Fig. 16. Goodness of fit χ^2/dof as a function of the timing offset parameter in units of the sample time.
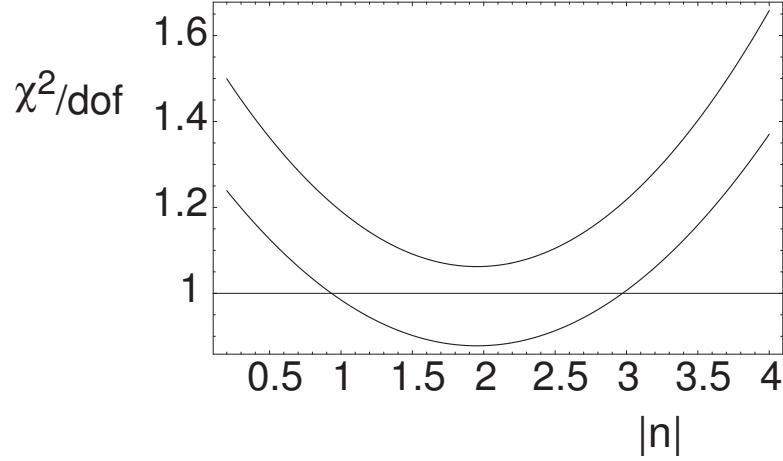

Fig. 17. Goodness of fit χ^2/dof as a function of the magnitude parameter $|n|$. Top curve uses $\alpha = 2.24$, bottom curve $\alpha = 2.44$. Both curves are acceptable due to degeneracy in estimating α (Fig. 15) . The statistical ideal of $\chi^2/dof = 1$ is the horizontal line.

The procedure is illustrated for Air Ch 1. The results of the other channels are comparable within uncertainties. Specifically, Air Ch 1 is the case best predicted, while the analysis of the other channels differ by less than a factor of 2 in χ^2 that can be attributed to the errors. The best fit to the timing offset parameter (Fig. 16) is about 120 sample time units. Recall that the filter bin dimension is $D = 32$ putting the onset around data point 1340. The timing offset is about 4-bin sizes of delay. The offset delay is consistent with model studies in which a few bins in a signal region are needed to develop significant filtered signal/noise. The best fit of the magnitude parameters is -1.95, indicating a polarity was reversed. A plot of χ^2/dof versus $|n|$ is shown in Fig. 17. The number of degrees of freedom dof are the number of points in the file (2460) minus the number of parameters (2). For these plots the estimated statistical fluctuation ("σ") in the denominator of χ^2 uses the

central value of the random walk parameter $\alpha = 2.24$ (Fig.15). The minimum χ^2/dof for the central value of α is about 1.06, which is perfectly acceptable.

Due to α, β parameter degeneracy, the random-walk parameter is not well fixed to the central value, and can vary without substantially changing the goodness of fit. From Fig. 15 the α parameter can be varied by about 50%, which suffices to account for variations of signal and noise observed between channels. A decrease of α by a mere 10% is statistically insignificant and causes the best fit χ^2 value to drop below 0.9 (Fig. 17, bottom curve). Assuming χ^2/dof should be near 1 on general statistical grounds, an error on parameter β can be assigned; the result is $\beta \sim 0.18 \rightarrow 0.21$, a relative change of order 15%, which stays well within the error ellipse of Fig. 15.

4   Summary

We conducted an experiment seeking to measure the impulsive fields from 120 GeV protons passing radio antennas in a sealed environment. A full system simulation was constructed describing charges moving at subliminal velocities in air (virtual radiation) and in wax (real Cherenkov radiation). One set of data simply measures propagation in air inside the tank. In the second dataset we observed signals consistent with real Cherenkov radiation in wax after beams passed through a lead preradiator. It is difficult to distinguish experimentally between the two experimental setups under the very near-field conditions of the apparatus. Both experiments registered signals in the causal onset region consistent with the virtual fields of charges moving at subluminal velocities. The signals demonstrate linear coherence and are consistent with simulation. We believe this is evidence for the first direct observation of radio frequency pulses from hadronic showers. The sizes of the two sets of signals were comparable and matched theoretical expectations.

A method extracting signal from data with very small signal-to-noise ratio has demonstrated novel and highly effective filtering techniques. The techniques are promising and may be applied broadly to improve experimental analysis of conditions with dramatically low signal to noise

References

[1] I. Kravchenko *et al.* (RICE collaboration), Astropart. Phys. **20**, 195 (2003) *ibid* **19**, 15 (2003); Phys. Rev. D **73**, 082002 (2006) [arXiv:astro-ph/0601148]; G. M. Frichter, J. P. Ralston and D. W. McKay, Phys. Rev. D **53**, 1684 (1996).

[2] See, e.g. P. Miocinovic *et al.* [The ANITA Collaboration], eConf **C**041213, 2516 (2004).

[3] F. Halzen, (arXiv:astro-ph/0602132); T. Montaruli [IceCube Collaboration], in *Proceedings of Vulcano Workshop 2006: Frontier Objects in Astrophysics and Particle Physics*, Vulcano, Italy, 22-27 (May 2006). (arXiv:astro-ph/0608140).

[4] G. Askarian, Sov.Phys.JETP 14, 441; D. Saltzberg *et al.*, Phys. Rev. Lett. **86**, 2802 (2001); P. W. Gorham, D. Saltzberg, R. C. Field, E. Guillian, R. Milincic, D. Walz and D. Williams, Phys. Rev. D **72**, 023002 (2005) [arXiv:astro-ph/0412128]; P. W. Gorham *et al.* [The ANITA Collaboration], Phys. Rev. Lett. **99**, 171101 (2007).

[5] A. Bean, R. V. Buniy and J. P. Ralston, AIP Conf. Proc. **579**, 234 (2001).

[6] O. Heaviside, *Electromagnetic Theory*, (Van Nostrand, 1912); L. Landau and L. Lifshitz, *Electrodynamics of Continuous Media*, Second Edition (Pergamon Press, 1984).

[7] The Fermilab test beam is summarized at http://www-ppd.fnal.gov/mtbf-w/

[8] S. Razzaque, S. Seunarine, S. W. Chambers, D. Z. Besson, D. W. McKay, J. P. Ralston and D. Seckel, Phys. Rev. D **69**, 047101 (2004) [arXiv:astro-ph/0306291]; Phys. Rev. D **69**, 047101 (2004) [arXiv:astro-ph/0306291].

[9] H. Karhunen, Ann. Acad. Science. Fenn, Ser A. I. 37, 1947; M. Loeve, supplement to P. Levy, *Processes Stochastic et Mouvement Brownien*, Paris, Gauthier Villars, 1948; H. Hotelling, J. Educ. Psychology 24, 417, 1933; ibid 24, 448, 1933; L. Scharfe, *Statistical Signal Processing*, (Wiley 1990).

[10] An interesting application of the Karhunen-Loeve procedure in cosmology, including cuts favoring subdominant subspaces, is given by R. Watkins, H. A. Feldman, S. W. Chambers, P. Gorman, and A. L. Melott, Ap. J. **564**, 534 (2002).

[11] Craig Moore, private communication.